

SEQUENCE LISTING



#4

(1) GENERAL INFORMATION:

- (i) APPLICANT: Ullrich, Axel
Aoki, Naohito
Kim, Yeong Woong
Wang, Hong Yang
Chen, Zhengjun
Naylor, Oliver
Kharitononkov, Alexei Igorevich
- (ii) TITLE OF INVENTION: NOVEL PTP20, PCP-2, BDP1, CLK,
AND SIRP POLYPEPTIDES AND RELATED
PRODUCTS AND METHODS
- (iii) NUMBER OF SEQUENCES: 38
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Lyon & Lyon
(B) STREET: 633 West Fifth Street
Suite 4700
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 90071-2066
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: IBM P.C. DOS 5.0
(D) SOFTWARE: FastSEQ for Windows 2.0
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: 08/877,150
(B) FILING DATE: June 17, 1997
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: U.S. 60/019,629
(B) FILING DATE: June 17, 1996
- (A) APPLICATION NUMBER: U.S. 60/023,485
(B) FILING DATE: August 9, 1996
- (A) APPLICATION NUMBER: U.S. 60/030,860
(B) FILING DATE: November 13, 1996

(A) APPLICATION NUMBER: U.S. 60/034,286
 (B) FILING DATE: December 19, 1996
 (A) APPLICATION NUMBER: U.S. 60/030,964
 (B) FILING DATE: November 15, 1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Warburg, Richard J.
 (B) REGISTRATION NUMBER: 32,327
 (C) REFERENCE/DOCKET NUMBER: 225/298

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (213) 489-1600
 (B) TELEFAX: (213) 955-0440
 (C) TELEX: 67-3510

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: "Xaa" in positions 3 and 5 stands
 for an unspecified amino acid.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Phe Trp Xaa Met Xaa Trp
 1 5

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: "Xaa" in position 6 stands for
 either Ser, Ile or Val.

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

His Cys Ser Ala Gly Xaa Gly
 1 5

- (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Phe Leu Glu Arg Leu Glu
 1 5

- (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (ix) FEATURE:

(D) OTHER INFORMATION: "Xaa" in positions 3 and 5 stands
 for an unspecified amino acid.

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Arg Trp Xaa Met Xaa Trp
 1 5

- (2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: "Xaa" in position 6 stands for either Ser, Ile or Val.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

His Cys Ser Ala Gly Xaa Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CTCTGTGTCC ACAGCAGTGC TGGCTGT

27

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

His Arg Asp Leu Ala Ala Arg
1 5

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: "Xaa" in position 2 stands for Val or Met. "Xaa" in position 5 stands for Tyr or Phe.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Asp Xaa Trp Ser Xaa Gly
 1 5

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CGGGATCCCT TCGCCTTGCA GCTTTGTC

28

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CGGAATTCCT AGACTGATAC AGTCTGTAAG

30

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Asp Leu Lys Pro Glu Asn
 1 5

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Ala Met Met Glu Arg Ile
1 5

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TATAGCGGCC GCTAGACTGA TACAGTCTGT

30

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TCCCCCGGGA TGCCCCATCC CCGAAGGTAC CA

32

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TATAGCGGCC GCTCACCGAC TGATATCCCG ACTGGAGTC

39

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TCCCCCGGGG AGACGATGCA TCACTGTAAG

30

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TATAGCGGCC GCGCTGGCCT GCACCTGTCA TCTGCTGGG

39

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CGGAATTCAT GCGGCATTCC AAACGAATC

30

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TATAGCGGCC GCCCTGACTC CCACTCATTT CCTTTTAA

39

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CGGAATTCCG CCACCATGGC CCCTATACTA GTTAT

36

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CCCAAGCTTG CCACCATGGC CCCTATACTA GTTAT

36

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GTAGCAGTAA GAATAGTTAA A

21

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GTTGCCCTGA GGATCATTA AAT

24

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GTTGCCCTGA GGATCATCCG GAAT

24

- (2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TACAATTCTC ACTGCTACAT GTAAGCCATC

30

- (2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Pro Ile Tyr Ser Phe Ile Gly Gly Glu His Phe Pro Arg
1 5 10

- (2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Ile Val Glu Pro Asp Thr Glu Ile Lys
1 5

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Tyr Gly Phe Ser Pro Arg
 1 5

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Ile Lys Glu Val Ala His Val Asn Leu Glu Val Arg
 1 5 10

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Val Ala Ala Gly Asp Ser Ala Thr
 1 5

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2226 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

```

GAATTCGGCGC ACGAGGCGGG TTGCAGTATG AGTCGCCAAT CGGACCTAGT GAGGAGCTTC      60
TTGGAGCAGC AGGAGGCCCG GGACCAACCG AAGGGGGCAA TCCTCGCCCG TGAGTTCAGC      120
GACATTAAAG CCGCTCTCAGT GCCTTGGAG ACTGAAGGTG TTGTCTCCAC TGAAGCCGCG      180
AGTCAGCAGG GAACTCTCAA GAGAACCGG TACAAGAGTG TGTATCGTGA TGATGAGCAGC      240
AGAGTTCATC TTTCCTCTGT CCAGGAGGAA GGACACGGAG ATTAACATTAA TGCCAATCTC      300
ATCCGGGGCA CAGATCGAAG CCAGGCCTAC ATTGGACGC AAGGACCCCT GCCTCACACT      360
CTGTTGGACT TCTGGCGCCT GGTTTGGGAG TTTGGAATCA AGGTGATCTT GATGGCCTGT      420
CAGGAGACAG AAAATGGACG GAGGAAGTGT GAACGCTACT GGGCCAGGA GCGGGAGCCT      480
CTACAGCGCG GGCCCTTTCTG CATCACCCCTG ACAAGGAGA CAGCAGTGAC TTCGGACATC      540
ACTCTCAGGA CCCTCCAGGT TACATTCCAG AAGGAATCCC GTCTGTGCA CCAGCTACAG      600
TACATGTCTT GGCCGGACCA CGGGGTTCCT AGCAGTTCGG ATCACATTCT CACCATGGTG      660
GAGGAGGCCG GTTGCTCTCA AGGACTTGGA CCTGGACCCC TCTGTGTCCA CTGCAGTGCT      720
GGCTGTGGAC GAACAGGTGT CTTGTGTGCT GTTGATTACG TGAGGCGAGT GCTTCTGACT      780
CAGACAATCC CACCCAATTT CAGCCTCTTT GAAGTGGTCC TGGAGATGCG GAAACAGCGA      840
CTTCGACGCG TGCACACAGA GGAGCAGTAC AGGTTCCTGT ACCACAGAT GGCTCAGTGA      900
TTCTCCGCGA CTCTCGACAA CACAGTCTCC CTCTACAGA ACCTCAAGGA GAACCGCGCT      960
CCAATCTGCA AGACTCTCTC GTCCCTCAGG ACCTCTCAG ACCTCCTGC CACATCCCG      1020
CCACTGGGTG GCCTTCTCAG GAGCATCTCG GTGCCGGGC CACCGACCTT TCCCATGGCT      1080
GACACTTACG CTGTGGTGCA GAAGCGTGGC GCTTCCGCA GCACAGGGCC GGGCACGCGG      1140
GGCCGCAACA GCACGGACAC CCCGACTTAC AGCCAGGTGG CTCACGTAAT CCAGCGGCGC      1200
GTGTCAACAC CCGAAACGCG GCAGGGGACA ACGGCAGTGG GCGGAGTTCC TGCGGATGAA      1260
AACCCTTCGCG GGCCGTGATGC CTATGAGGAA GTAACAGATG GAGCGCAGAG TGGTGGGCTA      1320
GGCTTCAACT TGCGCATTGG AAGACCTAAA GGGCCACGGG ATCCTCAGC GGAGTGGACA      1380
CGGCTGTAAAT GAGTGCTGTA CCACTTCCAG CCTGTCACTC AGTGGTGGCT GGGCGACTCG      1440
AACCCTCATG CTGCTGTGTG CTGTCTTATG TATGAGTGGG ACTCATGGGG TGTGAATCAA      1500
ATAAAGTTT CTACGGGTAG AAAAAACAA ATAGGGACTT TGGCCAGTGG TTATAGCAGT      1560
CAAAAGCCAG GGCTAGGAGG GGTAAAGTGG GGAGGTGGTG GATCTAGCTG GAGAAAGTTT      1620
AGGAAAGAGC ATCAAGAGTG AGCATCGCCA CTCTCTCTCC CATACACCTA CTGGAAAGTG      1680
CACCCAGAC AGAGTCCCTA CACTCAGAG CTGACAGTG CAGTCTGCTAC CTGGATGGAC      1740
ATGCTGCCCC TACAGCTCTA GACATGTCTA ATTGATCTCT CATGTAACT TCGAATGAGC      1800
TAGAAGATC TCCTGTCTGT CAGGGAAATG GATCACTAG TCAGTAAAT ACTGTCCAT      1860
CCAGAAGACA GAAGTGAAG ATACCGTCTT TCTCAAAATG GAAGAAATA GATCCTCAAG      1920
AATAAATGTA TGTACAATGC TCTACGCCCT GATCCTGCCC TGCCCTCACT CCATAATGTC      1980
ACAAACAAGT CAGGGTCTAT ATGACAGTTG TTCACTAGT CAGTCTCTAG TTGGCCCTCT      2040
CGAGGCTCAG ATAGTGCCTT CTGCAGACTC TTGAATGCC CGTCTTGAAC TTGATGAAAG      2100
CTTCTACCGG GAACCTTGA AATCATTAAT AATTATTAAT GTAGAATTCA ATAAAGAGTG      2160
GGTCAAAAC TCAAAAAA AAAAAA AAAAAA AAAAAA TCGAGAGTAC TTCTAGAGCG      2220
GCCGGG

```

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

```

Met Ser Arg Gln Ser Asp Leu Val Arg Ser Phe Leu Glu Gln Glu
 1             5             10             15
Ala Arg Asp His Arg Lys Gly Ala Ile Leu Ala Arg Glu Phe Ser Asp
 20             25             30

```

Ile Lys Ala Arg Ser Val Ala Trp Lys Thr Glu Gly Val Cys Ser Thr
 35 40 45
 Lys Ala Gly Ser Gln Gln Gly Asn Ser Lys Lys Asn Arg Tyr Lys Asp
 50 55 60
 Val Val Pro Tyr Asp Glu Thr Arg Val Ile Leu Ser Leu Leu Gln Glu
 65 70 75 80
 Glu Gly His Gly Asp Tyr Ile Asn Ala Asn Phe Ile Arg Gly Thr Asp
 85 90 95
 Gly Ser Gln Ala Tyr Ile Ala Thr Gln Gly Pro Leu Pro His Thr Leu
 100 105 110
 Leu Asp Phe Trp Arg Leu Val Trp Glu Phe Gly Ile Lys Val Ile Leu
 115 120 125
 Met Ala Cys Gln Glu Thr Glu Asn Gly Arg Arg Lys Cys Glu Arg Tyr
 130 135 140
 Trp Ala Gln Glu Arg Glu Pro Leu Gln Ala Gly Pro Phe Cys Ile Thr
 145 150 155 160
 Leu Thr Lys Glu Thr Ala Leu Thr Ser Asp Ile Thr Leu Arg Thr Leu
 165 170 175
 Gln Val Thr Phe Gln Lys Glu Ser Arg Pro Val His Gln Leu Gln Tyr
 180 185 190
 Met Ser Trp Pro Asp His Gly Val Pro Ser Ser Ser Asp His Ile Leu
 195 200 205
 Thr Met Val Glu Glu Ala Arg Cys Leu Gln Gly Leu Gly Pro Gly Pro
 210 215 220
 Leu Cys Val His Cys Ser Ala Gly Cys Gly Arg Thr Gly Val Leu Cys
 225 230 235 240
 Ala Val Asp Tyr Val Arg Gln Leu Leu Leu Thr Gln Thr Ile Pro Pro
 245 250 255
 Asn Phe Ser Leu Phe Glu Val Val Leu Glu Met Arg Lys Gln Arg Pro
 260 265 270
 Ala Ala Val Gln Thr Glu Glu Gln Tyr Arg Phe Leu Tyr His Thr Val
 275 280 285
 Ala Gln Leu Phe Ser Arg Thr Leu Gln Asn Asn Ser Pro Leu Tyr Gln
 290 295 300
 Asn Leu Lys Glu Asn Arg Ala Pro Ile Cys Lys Asp Ser Ser Ser Leu
 305 310 315 320
 Arg Thr Ser Ser Ala Leu Pro Ala Thr Ser Arg Pro Leu Gly Gly Val
 325 330 335
 Leu Arg Ser Ile Ser Val Pro Gly Pro Pro Thr Leu Pro Met Ala Asp
 340 345 350

Thr Tyr Ala Val Val Gln Lys Arg Gly Ala Ser Gly Ser Thr Gly Pro
 355 360 365

Gly Thr Arg Ala Pro Asn Ser Thr Asp Thr Pro Ile Tyr Ser Gln Val
 370 375 380

Ala Pro Arg Ile Gln Arg Pro Val Ser His Thr Glu Asn Ala Gln Gly
 385 390 395 400

Thr Thr Ala Leu Gly Arg Val Pro Ala Asp Glu Asn Pro Ser Gly Pro
 405 410 415

Asp Ala Tyr Glu Glu Val Thr Asp Gly Ala Gln Thr Gly Gly Leu Gly
 420 425 430

Phe Asn Leu Arg Ile Gly Arg Pro Lys Gly Pro Arg Asp Pro Pro Ala
 435 440 445

Glu Trp Thr Arg Val
 450

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5581 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

AATTCGGGGC	GCCAGTCCCG	CTCCGCGCGC	CGCCGCTCCG	CTCCGCGCTG	GGCTCCGGCT	60
CGCCTCGGGC	TGGGCTCGGG	CTCCGGGGGG	GGCGTCCCGG	CGCCGGGGCC	CGGAGACGGC	120
CGACCTCCAA	CCATGGCCCC	TGCCCAGGGG	CTCGTGCTGG	CACCTACCTT	CCAGCTCTGC	180
GCGCCGGAGA	CCGAGACTCC	GGCAGCTGGC	TGCACCTTCG	AGGAGGCCAG	TGACCCAGCA	240
GTGCCCTGCG	AGTACAGCCA	GGCCCACTAC	GATGACTTCC	AGTGGGAGCA	AGTGCGAATC	300
CACCTTGCCA	CCCGGGCACC	TGCGGACCTG	CCCCACGGCT	CCTACTTGAT	GGTCAACACT	360
TCCAGCATG	CCCCAGGCCA	CGAGCCCAT	GTCACTTTC	AGAGCCTGAG	CGAGAATGAT	420
ACCCACTGTG	TGCAGTTCAG	CTACTTCTCG	TACAGCGCGG	ACGACACAGG	CGGCACCTTG	480
CGCGTCTAGC	TGCGCGTTAA	TGGGGGCCCC	CTGGCGAGTG	CTGTGTGAA	TATGACTTGA	540
TCCCAAGGCC	GTCAGTGGCA	CCAGGCTGAG	CTGGCTGTCA	GCACCTTCTG	GCCCAATGAA	600
TATCAGGTGC	TGTTTGGAGC	CCTCATCTCC	CCAGACCCCA	GGGGCTACAT	GGGCTAGAT	660
GACATCTCTG	TTCTCAGCTA	CCCCTGCGCA	AAGGCCCCAC	ACTTCTCCCG	CCTGGCGGAC	720
GTGGAGGTCA	ACCGGGGCCA	GAAACGGTGC	TTCCAGTGGC	TGGCGCGGGG	AGAGCCCATG	780
CGCCAACGCT	TCTCTTGGCA	ACGGCAGAGC	GGGGCCCTGG	TGCCGGCCGG	GGCGTTCCGC	840
ACATCAGCCA	CGGGCTTCTT	GGCCACTTTC	CGCGTGCGTG	CCGTGAGCCG	CGCCGAGCAG	900
GACCTGTACC	GCTGTGTGTC	CCAGGCCCCG	CGCGCGGGCG	TCTCTAACTT	CCCGGAGCTC	960
ATCGTCAAGG	AGCCCCCAAC	TCCCATCGCG	CCCCCACAGC	TGCTGCGTGC	TGGCCCCACC	1020
TACCTCATCA	TCCAGCTCAA	CACCAACTCC	ATCATTGGCG	ACGGGCCGAT	CGTGCGCAAG	1080
GAGATTGAGT	ACCGCATGGC	GCGCGGGCCC	TGGGCTGAGG	TGCACGCCGT	CAGCCTGACG	1140
ACCTTACAAG	TGTGGCACTT	CGACCCCGAC	ACAGACTATG	AGATCAGCGT	GCTGCTCAGG	1200
CTTCCCGGAG	ACCGCGGCAC	TGGCGCTGGG	GCCACCCCTC	ATCAGCCGCA	CCAAATGGCG	1260
AGAGCCCATG	AGGGCCCCAA	AGGCTCTGAG	TTTGCTGAGA	TCCAGGCCCG	TCAAGTGACC	1320
CTGCAGTGGG	AACCACTGGG	TACACACGTG	ACGCGTGGCC	ACACCTATAC	TGTGCTCTGT	1380
TGCTATCACT	ACACCTTGGG	CAGCAGCCAC	AACCAAGACCA	TCCGAGAGTG	TGTGAGAGCA	1440
GAGCAAGGTG	TCAGCGGCTA	CACCATCAAG	AACCTGCTCG	CCTATCGGAA	CGTCCAGCTG	1500
AGGCTTGTCC	TCACTAACCC	TGAGGGGGCG	AAGAGGGGCA	AGGAGGTCAC	TTTCCAGAGC	1560
GATGAGGATG	TGCCAGTGG	GATTGCAGCC	GAGTCCCTGA	CCTTCACTCC	ACTGGAGGAC	1620

ATGATCTTCC	TCAAGTGGGA	GGAGCCCCG	GAGGCCAATG	GTCTCATCAC	CCAGTATGAG	1680
ATCAGCTTACC	AGAGCATCGA	GTCAATCAGAC	CGGGCAGTGA	ACGTGCGCAGG	CCCACGACGT	1740
ACCATCTTCCA	ACTCTGGGGA	TGAGACCTTAC	CATGTCTTCT	CCAACCTTGCA	CCGACGGCACC	1800
ACCTACCTGT	TCTCCGTGGC	TGCCTCGACA	GATCAAGGCT	TCGGCCAGGCT	CGCACTCACT	1860
GAGATAACCA	CTAACACTCT	TGCTCCCAAG	TTTGATTTATG	CGCACTATGCT	CGGTCGCGCT	1920
GGCAGGCTCTG	AGAACACCAT	CACCGTGTCTG	TGAGGCGCGG	CACAGGGCGG	CGGTGCGGCC	1980
ATCAGTGTGT	ACCAGGTGAT	TGTGGAGGAG	GAGCGGGCGG	GAGGCTGCGG	CGGACGAGG	2040
TGGACAGGAC	TGCTTCCGAG	TGCCATTGAC	CTTCTGAGCG	GGCTTGGCCC	CAGGCTGGTG	2100
CACCTACTTCT	GGGCGGAAGT	GGCGGCCAGC	AGTCTACTGT	AGGCCATGCC	CTTTACCGTG	2160
GGTGACACAC	AGACCTTACG	AGGCTTCTGG	AACCCACCAC	TTGAGCTTAG	GAAAGGCTAT	2220
CTCATCTACT	TCCAGGACAGC	AAGCCACCTG	AAGGGGGAGA	CCCGGCTGAA	TTGCATCCCG	2280
ATTGCCAGGA	AACTGTCTCTG	CAAGGAAAGC	AAGCGGCCCT	TGGAGGTGTC	CCAGAGATCG	2340
GAGGAGATGG	GGCTTATCTCT	GGGCATCTGT	GCAGGGGGGG	TTGCTGTCTT	CATCTTCTCT	2400
CTGGGTGCCA	TCAATTGTCTAT	CATCCGCAAA	GGGAAGCCGG	TGAACATGAC	CAAGGGCCACC	2460
GTCAACTTACC	GGCAGAGGAA	GACACACATG	ATCAGCGCCC	TGGACCGCAG	CTTACAGAGC	2520
CAGAGCACAC	TCCAGAGAGGA	CGAGCGGCTG	GGCTGTCTCT	TCATGGACAGT	CCATGGCTTAC	2580
AGACCCGGGG	GAGACCCAGC	CAGCGGTGGG	GTCACTGAGG	CCAGACGAGCT	CCTGGGGGGC	2640
TCCCCGAGGC	GTCTCTGTGG	CGGGAAGGGC	TCCCACTTAC	ACACGGGACA	GCTCAAGCCT	2700
CGCGTGGCTG	CTCGAGACCT	TCTCGACAC	ATCAACAGAG	TGAAGACGGG	CGAGGGTTAC	2760
GGCTTCAAGC	AGGAGTATGA	GAGCTTCTTT	GAAAGCTGGG	ACGCCACAAA	GAAAGAAAGC	2820
AAGGTCAAGG	GCAGCCGGCA	GGAGCCAATG	GCTGCCATAT	ATCGGCACCG	AGTGAACACT	2880
CACCCGATGC	TGGGAGACCC	CAATGCCGAG	TACATTAATG	CCAACCTACAT	AGATGGTTAC	2940
CACAGGTCAA	ACCACTTTCAT	AGCCACTTCA	GGGCGGAAG	CTGAGATGGT	CTATGACTTCT	3000
TGGGTGATCT	TGTGGCAGGA	GCACTGTTCC	AGCATCGTCA	TGATACACAA	GCTGGGTCAG	3060
TGGGCGAGGG	TGAAATCTCT	ACGGTACTGG	CGGAGGAGCT	CAGACACCTA	CGGGGACATC	3120
AGAATTTATG	TGTTGAAGAC	AGAGACCCCTG	GCTGAGTATG	TGCTGCGCAC	TTTTGCCCTG	3180
GAGCGGAGAG	GCTACTTCTG	CGGCGACGAG	GTCCCGCATG	CCCACTTCAC	GCGGTGCCCA	3240
GAGCATGGGG	TCCCTTACCA	TGCCACGGGG	GTCTGGCTTT	TCACTCCGGG	GGTGAAGGCC	3300
TCCACCCCACT	CTGAGTGGCG	GGCCATTTGTC	ATGCCACTGCA	GGCGGGGACG	CGGGCGACCA	3360
CTGTGCTATA	TGCTGTGGGA	TGTGATGCTG	GACATGCGAG	AGTGTGAGGG	CGTGTGGAGC	3420
ATTTACAACCT	GTGTGAAGAC	CGCGGTGTCA	ACATGATCCA	GACTGAGGAG	GACTGAGGAG	3480
CAGTACACTCT	CTATTACTGA	TGCAATCTCTG	GAGGCCGTGC	TGTGTGGGGA	GACCACTATC	3540
CCCTGTGATCT	AGTTCAAGGC	CACCTACAAG	GAGATGATCC	GCATTGATTC	TCAGAGTAAT	3600
TCTTCCACAG	TGCGGGAAGA	GTTTCCAGAC	CTGAACTCGG	TCACCCCGCC	GCTGGACGTG	3660
GAGGAGTGCA	GCATCGCCTT	GTTGCCCCGG	AACCGCGACA	AGAACCAGCAG	CATGGACGTC	3720
CTGCGCGCGC	ACCGCTGCCT	GCCTTCTCTT	ATCTCCAAGT	ATGGGGACTC	CAACAACATC	3780
ATTAAATGAG	CCCTGACGTA	CAGCTACACA	CGGAGGTGG	CCTTCATGTT	GACCTGTGAC	3840
CGCGTGCAGA	GCACACGCTG	CGACTTCTGG	CGGCTGTGCT	ACGATTACGG	GTGCACTCTC	3900
ATCGTCAATC	TCAACAGAGT	GAACCAAGTCC	AACTCCGCTT	GGCCTTGCTT	GCAGTACTGG	3960
CCAGAGCCAG	CGCGGAGCA	ATATGGCCTC	ATGGAGGTGG	AGTTTATGTC	ATGGAGGCTC	4020
GATGAAGACT	TAGTGGCTCG	AGTCTTCCGG	GTGCAGAAC	TCTCTGGGTT	GCAGGAGGGA	4080
GACCTGTCTG	TGCGGCACTT	CCAGTTCTCT	CGCTGGTCTG	CATACCGGGA	CACACCTGAC	4140
TCCAGGAAGG	CTTCTTGGCA	CTGTGTGGCT	GAGGTGGAGA	AGTGGCAGCG	CGAGAGTGGG	4200
GATGGCGCA	CTGCTTGCA	CTGCCTAAT	CGGCGAGGAC	CTGCGGACAC	CTTCTGCGCC	4260
TGCCCAACGG	TCTTGGAGAT	GATCCGCTGC	CACACTCTGT	TGAGCTTTTG	CTTCTTCTGC	4320
CAAAACCTCT	GGAATCAACA	ACCCAACATG	GTGAGAGACC	TGGATCAGTA	CCACTTTTGC	4380
TACAGTGTGG	CTGGAGAGTA	CTTGGAGGGG	CTGAGCTCAA	GATAGCGGCT	CCCTGGCTCT	4440
GGGACCCACG	TGCACACTCA	GGGCGAGACC	CACCATCTCT	GACTGGGCGG	GAAATTCAGT	4500
GCCTCTTGCT	CTCCCAAAAC	ACACTCCCAT	GGGGCAAGCA	CTGAGTGGGA	TGCTGGGCTA	4560
TCTTGTCTCC	CCCTTCACTG	TGGGCAAGGG	CTTTCGCTGT	TCCCATGGGC	GGGTGGTGGG	4620
CAAAGGAGGA	GTTTGACAAG	CTTGACACCC	ACCCCACTCT	CCATAGGGTC	GCTGAGGCCCT	4680
GTGCTGAGAG	GCCTGGTGTCT	GCCTGGCAGA	GTGACAAAGG	CTCAGGACGG	CTGGCTCTGG	4740
GGGACTCAGG	CCAAGGGGTT	TGGCAGGATC	CTGGGTTTTG	GGAGGGATGA	GTGAGGCCCT	4800
GCAGAGGAGCA	TCCCAAGGCC	AGGTTCCCA	CTAGCCTGCC	CCCTCTGCAT	GTGGGTATAG	4860
GATCTACTGG	GACTTGGCAT	TTAGGATTCC	ATCTGGGGGA	CCCCCTGAAG	GTCCCCCAAC	4920
AGCAGTCTCT	AATTCTGATA	GCGAGTGGGG	CACACTGACT	GCTCTTCCCA	GGGGAAGTGC	4980
AGCGCCCTCT	TCCCAATGTC	CCCTCTCAGC	CCCTGAGATA	TTTGTGCTCA	TATCTCTCCC	5040
CACCTGTCTTC	CAGCTTGATCA	CTGCTGATCA	CTGCTGATCA	GACTCTGCTT	CTACTCTGCT	5100
TCCCAATGGGG	GGCTCTTCTC	CTGCTTGACC	CACGTGTGCA	GACTGATGCA	GAATGAAGCT	5160
CCTCTTCTCT	TAAATCTCAG	GCCTCACTGG	CCTGCTCTGC	TCAGTCTGGG	CCAGTGACAA	5220
TCTGCAAGGC	TGAACAACAG	CCCTTGGGGT	TGAGGCCCTT	GTGGCTCTCT	GTCAAGCTGC	5280
CCGTGGTGGG	GAGGGGCGAGT	GTTAGAGCAG	GGCTGTGTCAT	ACCTCTTGGA	GTTCAAGACA	5340

AGAGGTAGGA	CCAGTGCTTT	TTTGTTCTT	TTGTTATTT	TGGTGGGTG	GGTGGGAAGG	5400
TCTCTTTAAA	ATGGGGCAGG	CCACACCCCC	ATTCCGTGCC	TCAATTTCCT	CATCTGTAAA	5460
CTGTAGATAT	GACTACTGAC	CTACCTCGCA	GGGGGCTGTG	GGGAGGCATA	AGCTGATGTT	5520
TGTAAGCGC	TTTGTAATA	AACGTGCTCT	CTGAATGCCA	AAAAAAAAAA	AACAAAAAAA	5580
A						5581

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1430 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Met	Ala	Arg	Ala	Gln	Ala	Leu	Val	Leu	Ala	Leu	Thr	Phe	Gln	Leu	Cys
1				5				10						15	
Ala	Pro	Glu	Thr	Glu	Thr	Pro	Ala	Ala	Gly	Cys	Thr	Phe	Glu	Glu	Ala
	20						25						30		
Ser	Asp	Pro	Ala	Val	Pro	Cys	Glu	Tyr	Ser	Gln	Ala	Gln	Tyr	Asp	Asp
	35					40				45					
Phe	Gln	Trp	Glu	Gln	Val	Arg	Ile	His	Pro	Gly	Thr	Arg	Ala	Pro	Ala
	50				55					60					
Asp	Leu	Pro	His	Gly	Ser	Tyr	Leu	Met	Val	Asn	Thr	Ser	Gln	His	Ala
	65				70				75					80	
Pro	Gly	Gln	Arg	Ala	His	Val	Ile	Phe	Gln	Ser	Leu	Ser	Glu	Asn	Asp
		85						90					95		
Thr	His	Cys	Val	Gln	Phe	Ser	Tyr	Phe	Leu	Tyr	Ser	Arg	Asp	Gly	Thr
		100						105					110		
Gly	Gly	Thr	Leu	Arg	Val	Tyr	Val	Arg	Val	Asn	Gly	Gly	Pro	Leu	Ala
	115						120					125			
Ser	Ala	Val	Trp	Asn	Met	Thr	Gly	Ser	His	Gly	Arg	Gln	Trp	His	Gln
	130				135						140				
Ala	Glu	Leu	Ala	Val	Ser	Thr	Phe	Trp	Pro	Asn	Glu	Tyr	Gln	Val	Leu
	145				150				155					160	
Phe	Glu	Ala	Leu	Ile	Ser	Pro	Asp	Arg	Arg	Gly	Tyr	Met	Gly	Leu	Asp
		165					170						175		
Asp	Ile	Leu	Leu	Leu	Ser	Tyr	Pro	Cys	Ala	Lys	Ala	Pro	His	Phe	Ser
		180					185						190		
Arg	Leu	Gly	Asp	Val	Glu	Val	Asn	Ala	Gly	Gln	Asn	Ala	Ser	Phe	Gln
	195						200					205			
Cys	Met	Ala	Ala	Gly	Glu	Pro	Met	Arg	Gln	Arg	Phe	Leu	Leu	Gln	Arg
	210						215					220			

Gln Ser Gly Ala Leu Val Pro Ala Gly Ala Phe Gly Thr Ser Ala Thr
 225 230 235 240
 Gly Phe Leu Ala Thr Phe Pro Leu Ala Ala Val Ser Arg Ala Glu Gln
 245 250 255
 Asp Leu Tyr Arg Cys Val Ser Gln Ala Pro Arg Gly Gly Val Ser Asn
 260 265 270
 Phe Pro Glu Leu Ile Val Lys Glu Pro Pro Thr Pro Ile Ala Pro Pro
 275 280 285
 Gln Leu Leu Arg Ala Gly Pro Thr Tyr Leu Ile Ile Gln Leu Asn Thr
 290 295 300
 Asn Ser Ile Ile Gly Asp Gly Pro Ile Val Arg Lys Glu Ile Glu Tyr
 305 310 315 320
 Arg Met Ala Arg Gly Pro Trp Ala Glu Val His Ala Val Ser Leu Gln
 325 330 335
 Thr Tyr Lys Leu Trp His Leu Asp Pro Asp Thr Asp Tyr Glu Ile Ser
 340 345 350
 Val Leu Leu Thr Arg Pro Gly Asp Gly Gly Thr Gly Arg Trp Ala Thr
 355 360 365
 Pro His Gln Pro His Gln Met Arg Arg Ala His Glu Gly Pro Lys Gly
 370 375 380
 Leu Ala Phe Ala Glu Ile Gln Ala Arg Gln Leu Thr Leu Gln Trp Glu
 385 390 395 400
 Pro Leu Gly Tyr Asn Val Thr Arg Cys His Thr Tyr Thr Val Ser Leu
 405 410 415
 Cys Tyr His Tyr Thr Leu Gly Ser Ser His Asn Gln Thr Ile Arg Glu
 420 425 430
 Cys Val Lys Thr Glu Gln Gly Val Ser Arg Tyr Thr Ile Lys Asn Leu
 435 440 445
 Leu Pro Tyr Arg Asn Val His Val Arg Leu Val Leu Thr Asn Pro Glu
 450 455 460
 Gly Arg Lys Glu Gly Lys Glu Val Thr Phe Gln Thr Asp Glu Asp Val
 465 470 475 480
 Pro Ser Gly Ile Ala Ala Glu Ser Leu Thr Phe Thr Pro Leu Glu Asp
 485 490 495
 Met Ile Phe Leu Lys Trp Glu Glu Pro Gln Glu Pro Asn Gly Leu Ile
 500 505 510
 Thr Gln Tyr Glu Ile Ser Tyr Gln Ser Ile Glu Ser Ser Asp Pro Ala
 515 520 525
 Val Asn Val Pro Gly Pro Arg Arg Thr Ile Ser Lys Leu Arg Asn Glu
 530 535 540

Thr Tyr His Val Phe Ser Asn Leu His Pro Gly Thr Thr Tyr Leu Phe
 545 550 555 560
 Ser Val Arg Ala Arg Thr Gly Lys Gly Phe Gly Gln Ala Ala Leu Thr
 565 570 575
 Glu Ile Thr Thr Asn Ile Ser Ala Pro Ser Phe Asp Tyr Ala Asp Met
 580 585 590
 Pro Ser Pro Leu Gly Glu Ser Glu Asn Thr Ile Thr Val Leu Leu Arg
 595 600 605
 Pro Ala Gln Gly Arg Gly Ala Pro Ile Ser Val Tyr Gln Val Ile Val
 610 615 620
 Glu Glu Glu Arg Ala Arg Gly Cys Gly Gly Thr Arg Trp Thr Gly Leu
 625 630 635 640
 Leu Pro Ser Ala Ile Asp Leu Arg Gly Gly Ala Gly Pro Arg Leu Val
 645 650 655
 His Tyr Phe Gly Ala Glu Leu Ala Ala Ser Ser Leu Pro Glu Ala Met
 660 665 670
 Pro Phe Thr Val Gly Asp Asn Gln Thr Tyr Arg Gly Phe Trp Asn Pro
 675 680 685
 Pro Leu Glu Pro Arg Lys Ala Tyr Leu Ile Tyr Phe Gln Ala Ala Ser
 690 695 700
 His Leu Lys Gly Glu Thr Arg Leu Asn Cys Ile Arg Ile Ala Arg Lys
 705 710 715 720
 Ala Ala Cys Lys Glu Ser Lys Arg Pro Leu Glu Val Ser Gln Arg Ser
 725 730 735
 Glu Glu Met Gly Leu Ile Leu Gly Ile Cys Ala Gly Gly Leu Ala Val
 740 745 750
 Leu Ile Leu Leu Leu Gly Ala Ile Ile Val Ile Ile Arg Lys Gly Lys
 755 760 765
 Pro Val Asn Met Thr Lys Ala Thr Val Asn Tyr Arg Gln Glu Lys Thr
 770 775 780
 His Met Ile Ser Ala Val Asp Arg Ser Phe Thr Asp Gln Ser Thr Leu
 785 790 795 800
 Gln Glu Asp Glu Arg Leu Gly Leu Ser Phe Met Asp Thr His Gly Tyr
 805 810 815
 Ser Thr Arg Gly Asp Gln Arg Ser Gly Gly Val Thr Glu Ala Ser Ser
 820 825 830
 Leu Leu Gly Gly Ser Pro Arg Arg Pro Cys Gly Arg Lys Gly Ser Pro
 835 840 845
 Tyr His Thr Gly Gln Leu His Pro Ala Val Arg Val Ala Asp Leu Leu
 850 855 860
 Gln His Ile Asn Gln Met Lys Thr Ala Glu Gly Tyr Gly Phe Lys Gln
 865 870 875 880

Glu Tyr Glu Ser Phe Phe Glu Gly Trp Asp Ala Thr Lys Lys Lys Asp
 885 890 895
 Lys Val Lys Gly Ser Arg Gln Glu Pro Met Pro Ala Tyr Asp Arg His
 900 905 910
 Arg Val Lys Leu His Pro Met Leu Gly Asp Pro Asn Ala Asp Tyr Ile
 915 920 925
 Asn Ala Asn Tyr Ile Asp Gly Tyr His Arg Ser Asn His Phe Ile Ala
 930 935 940
 Thr Gln Gly Pro Lys Pro Glu Met Val Tyr Asp Phe Trp Arg Met Val
 945 950 955 960
 Trp Gln Glu His Cys Ser Ser Ile Val Met Ile Thr Lys Leu Val Glu
 965 970 975
 Val Gly Arg Val Lys Cys Ser Arg Tyr Trp Pro Glu Asp Ser Asp Thr
 980 985 990
 Tyr Gly Asp Ile Lys Ile Met Leu Val Lys Thr Glu Thr Leu Ala Glu
 995 1000 1005
 Tyr Val Val Arg Thr Phe Ala Leu Glu Arg Arg Gly Tyr Ser Ala Arg
 1010 1015 1020
 His Glu Val Arg Gln Ser His Phe Thr Ala Trp Pro Glu His Gly Val
 1025 1030 1035 1040
 Pro Tyr His Ala Thr Gly Leu Leu Ala Phe Ile Arg Arg Val Lys Ala
 1045 1050 1055
 Ser Thr Pro Pro Asp Ala Gly Pro Ile Val Ile His Cys Ser Ala Gly
 1060 1065 1070
 Thr Gly Arg Thr Arg Cys Tyr Ile Val Leu Asp Val Met Leu Asp Met
 1075 1080 1085
 Ala Glu Cys Glu Gly Val Val Asp Ile Tyr Asn Cys Val Lys Thr Leu
 1090 1095 1100
 Cys Ser Arg Arg Val Asn Met Ile Gln Thr Glu Glu Gln Tyr Ile Phe
 1105 1110 1115 1120
 Ile His Asp Ala Ile Leu Glu Ala Cys Leu Cys Gly Glu Thr Thr Ile
 1125 1130 1135
 Pro Val Ser Glu Phe Lys Ala Thr Tyr Lys Glu Met Ile Arg Ile Asp
 1140 1145 1150
 Pro Gln Ser Asn Ser Ser Gln Leu Arg Glu Glu Phe Gln Thr Leu Asn
 1155 1160 1165
 Ser Val Thr Pro Pro Leu Asp Val Glu Glu Cys Ser Ile Ala Leu Leu
 1170 1175 1180
 Pro Arg Asn Arg Asp Lys Asn Arg Ser Met Asp Val Leu Pro Pro Asp
 1185 1190 1195 1200

Arg Cys Leu Pro Phe Leu Ile Ser Thr Asp Gly Asp Ser Asn Asn Tyr
 1205 1210 1215
 Ile Asn Ala Ala Leu Thr Asp Ser Tyr Thr Arg Arg Ser Ala Phe Met
 1220 1225 1230
 Val Thr Leu His Pro Leu Gln Ser Thr Thr Pro Asp Phe Trp Arg Leu
 1235 1240 1245
 Val Tyr Asp Tyr Gly Cys Thr Ser Ile Val Met Leu Asn Gln Leu Asn
 1250 1255 1260
 Gln Ser Asn Ser Ala Trp Pro Cys Leu Gln Tyr Trp Pro Glu Pro Gly
 1265 1270 1275 1280
 Arg Gln Gln Tyr Gly Leu Met Glu Val Glu Phe Met Ser Gly Thr Ala
 1285 1290 1295
 Asp Glu Asp Leu Val Ala Arg Val Phe Arg Val Gln Asn Ile Ser Arg
 1300 1305 1310
 Leu Gln Glu Gly Asp Leu Leu Val Arg His Phe Gln Phe Leu Arg Trp
 1315 1320 1325
 Ser Ala Tyr Arg Asp Thr Pro Asp Ser Lys Lys Ala Phe Leu His Leu
 1330 1335 1340
 Leu Ala Glu Val Asp Lys Trp Gln Ala Glu Ser Gly Asp Gly Arg Thr
 1345 1350 1355 1360
 Ile Val His Cys Leu Asn Gly Gly Gly Arg Ser Gly Thr Phe Cys Ala
 1365 1370 1375
 Cys Ala Thr Val Leu Glu Met Ile Arg Cys His Asn Leu Val Asp Val
 1380 1385 1390
 Phe Phe Ala Ala Gln Thr Leu Arg Asn Tyr Lys Pro Asn Met Val Glu
 1395 1400 1405
 Thr Met Asp Gln Tyr His Phe Cys Tyr Asp Val Ala Leu Glu Tyr Leu
 1410 1415 1420
 Glu Gly Leu Glu Ser Arg
 1425 1430

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2810 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

GAATTCGGCA CGAGCGGGCT GGACCTTCT	CGCCCGCGGC	GCCATGAGCC	CGAGCCTGGA	60
CTCGGCGCGG AGCTTCCTGG ACCGGCTGGA	AGCGCGGGGC	GGCGGGGAGG	GGGCAGTCCT	120
CGCCGCGCAG TTCAGCGACA TCCAGGCGCTG	CTCGGCGCGC	TGGAAGGCTG	ACGGCGGTGT	180

CTCCACCGTG	CGCCGCGAGT	GGCCAGAGAA	CGTGAGGAAG	AACCGCTACA	AAGACGTGCT	240
GCCTTAATGAT	CAGACGGGAG	TAATCTCTCT	CCTGCTCCAG	GAAGAGGGAC	ACAGCGGACTA	300
CATTAAATGGC	AACCTFCATCC	GGGGCGTGGA	TGGAAGCCTG	GCCTACATTG	CCAGCGCAAGG	360
ACCCTTGCTCT	CACACCCCTGC	TAGACTTCTG	GAGACTGGTC	TGGGAGTTTG	GGGTCAAGGT	420
GATCCTGTATG	GCCTGTGCGAG	AGATAGAGAA	TGGGCGGAAA	AGGTGTGAGC	GGTACTGGGC	480
CCAGGAGGAC	GAGGCCACTGC	AGACTGGGCT	TTTCTGCATC	ACTCTGATAA	AGGAGAAGTG	540
GCTGAARTAG	GACATCATGCT	TCAGGACCCCT	CAAGGTACCA	TTCCAGAAAG	AGTCCCGTTC	600
TGTTGTACAG	CTACAGTATA	TGTCCTTGCC	AGACCGTGGG	GTCCCCAGCA	GTCTGAGACA	660
CATGCTCGCC	ATGGTGGAGG	AAGCCGCTG	CTCCAGGGA	TCTGGCCCTG	AACCCCTCTG	720
TGTCGACTGT	AGTGGCGGTT	TGGGGCGAAC	AGGGCTCTTG	TGACCCGTGG	ATTATGTGAG	780
GCAGCTGCTC	CTGACCCAGA	TGATCCCAAC	TGACTTCAGT	CTCTTTGATG	TGCTCCTTAA	840
GATGAGGAAG	CAGCGGCTG	CGGCGGTGCA	GACAGAGGAG	CAGTACAGGT	TCCTGTACCA	900
CACGGTGGCT	CAGATGTTCT	GCTCCCACT	CCAGAATGCC	AGCCCCCACT	ACCAGAACAT	960
CAAGAGAAAT	TGTGCCCCAC	TCTACGACGA	TGCCCTTCTC	CTCCGGACTC	CCCAGGCACT	1020
TCTCGCCATA	CCCCGCCAC	CAGGAGGGGT	CCTCAGGAGC	ATCTCTGTGC	CCGGGTCCCC	1080
GGGCCACGCC	ATGGCTGACA	CCTACGCGGA	GGAGCAGAAG	CGCGGGGCTC	CAGGGGCGCT	1140
CGGGAGTGGG	ACGCGACGCG	GGACGGGGAC	GGGGCGCGCG	AGCGCGGAGG	AGGGCGCGCT	1200
CTACAGCAAG	GTGACGCGCG	CGGCCACGG	ACCCGGGGCG	CACGCGGAGG	ACGCGAGGGG	1260
GACGCTGCCT	GGCCGCGTTC	CTGCTGACCA	AAGTCTTGCC	GGATCTGGCG	CCTACGAGGA	1320
CGTGGCGGGT	GGAGCTCAGA	CCGGTGGGCT	AGGTTTCAAC	CTGCGCATTG	GGAGGCCGAA	1380
GGGTCCCCGG	GACCCCGCTG	CTGAGTGGAC	CCGGGTGTAA	CTCTACCGCT	AGTTCTTGCC	1440
TGTTGCTCTT	TGTGAGCTCG	GACTGCTGAT	CGCCCGTGTC	TGCTGAGCGC	CGTGCCGAGA	1500
ATGGAACAGC	TGGGCTTGA	TCAAAAGTTA	AGTTTCTCAG	GGTGGGAAT	GTGGGGGCTT	1560
TGCCCAATGA	CTGTAGCATT	CAAGGCTTGA	GGCTGGAGGA	GGTAGCTAGG	GTATATGTGC	1620
TGTTGAGGCT	GACACAGACA	GATTCAGAAA	AGAAGATCAG	GAAAGGCGAT	AGCCCTGTAG	1680
TATTGAAGGG	GAGAAGGGAC	AGATGAGCTT	CCGGAGACTG	CTCTCTCTAC	CACACAGCAC	1740
TAGTCTCATC	TGAGCACTG	AGCCTCCCTC	ACTTGGACAC	TCAGGGGACC	ACACAGAGAA	1800
TGTGATGAGC	ACTTGGCCAT	CCAGGCGAGAA	CTAAGCCAGG	CATAACCA	GCCAAGCAGA	1860
TTAAACCCAG	GCAGACCGAT	AAAAGAGACT	CCAGATAGGC	AGACAGACAG	AGTCCACACC	1920
AACCTGGACA	GACAGCCAAA	GCTTCAGAGA	TACAGTCCAC	AGGTGGACAA	AGGATCCCCC	1980
AGCCAGAGAG	AGAGAGACCA	GCCAACAGCT	TGATAGACCA	TGTCAGCCAG	AGAGACCAACC	2040
AAACACAGCC	CCCCAAAGAC	AGACATCTCT	GCTAGCTGGA	CAGCCAGGTT	GACCCCTCTAA	2100
GTTAGTCAGA	TTACTAGACA	GATATAAACA	GATCCCTGCG	TGAACAGATA	TACAGAGTTC	2160
TCAGACCCCA	CTCCCTCAGG	TGGGCTGGCT	GGCTGACAGA	CTTCTTGCC	AGACGACTC	2220
CTAACCAACC	AGATGGACTG	CCAGACAGGC	AGACATAGT	CCACATGGA	TCCTGACATC	2280
CCAGCCAGCC	GGCCAGACTG	TCATCTTGAT	GTCTGATGG	ATGGACCCCA	GCTAGTCAGA	2340
CATGATCTTC	CAGATTGACA	GACAAGTCCC	CCAATAGAT	ACACATCTCC	AGCTATTTCAG	2400
ACAGATGAG	CCCCAGCAA	TCAGGACCTA	TCTAGGCAGA	CCCCCGCCA	ACCCCGGCCA	2460
GACGAGCTCC	CAACCAAGACT	GACCCCTTGC	TGTTACACA	GCCTGCGGAG	TAGCTGGGAC	2520
TACAGTCTTA	ATTTTITTTT	TTTTTAAAGAA	ATGAGTITTT	GCCATGTTGC	CCAGACTGGT	2580
CTTGAACCTC	CAACCTCAAG	CAATCTCTCT	GCCTCAGCCT	CCCAAAGTGC	TGAGATTACA	2640
GGTGTGAGTC	ACAGGCTCA	GCCCCCTAAG	ATTTGAAACA	CTTTAAATGG	CCCATGGTAG	2700
GTTCTCTGCT	AGGATAAAC	ATTAAGTGGC	TGTTAAAGAA	AATAAAAGGA	GGACACGCTCT	2760
CGTGTCAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA		2810

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 458 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Met Ser Arg Ser Leu Asp Ser Ala Arg Ser Phe Leu Glu Arg Leu Glu
 1 5 10 15
 Ala Arg Gly Gly Arg Glu Gly Ala Val Leu Ala Gly Glu Phe Ser Asp
 20 25 30
 Ile Gln Ala Cys Ser Ala Ala Trp Lys Ala Asp Gly Val Cys Ser Thr
 35 40 45
 Val Ala Gly Ser Arg Pro Glu Asn Val Arg Lys Asn Arg Tyr Lys Asp
 50 55 60
 Val Leu Pro Tyr Asp Gln Thr Arg Val Ile Leu Ser Leu Leu Gln Glu
 65 70 75 80
 Glu Gly His Ser Asp Tyr Ile Asn Gly Asn Phe Ile Arg Gly Val Asp
 85 90 95
 Gly Ser Leu Ala Tyr Ile Ala Thr Gln Gly Pro Leu Pro His Thr Leu
 100 105 110
 Leu Asp Phe Trp Arg Leu Val Trp Glu Phe Gly Val Lys Val Ile Leu
 115 120 125
 Met Ala Cys Arg Glu Ile Glu Asn Gly Arg Lys Arg Cys Glu Arg Tyr
 130 135 140
 Trp Ala Gln Glu Gln Glu Pro Leu Gln Thr Gly Leu Phe Cys Ile Thr
 145 150 155 160
 Leu Ile Lys Glu Lys Trp Leu Asn Glu Asp Ile Met Leu Arg Thr Leu
 165 170 175
 Lys Val Thr Phe Gln Lys Glu Ser Arg Ser Val Tyr Gln Leu Gln Tyr
 180 185 190
 Met Ser Trp Pro Asp Arg Gly Val Pro Ser Ser Pro Asp His Met Leu
 195 200 205
 Ala Met Val Glu Glu Ala Arg Arg Leu Gln Gly Ser Gly Pro Glu Pro
 210 215 220
 Leu Cys Val His Cys Ser Ala Gly Cys Gly Arg Thr Gly Val Leu Cys
 225 230 235 240
 Thr Val Asp Tyr Val Arg Gln Leu Leu Leu Thr Gln Met Ile Pro Pro
 245 250 255
 Asp Phe Ser Leu Phe Asp Val Val Leu Lys Met Arg Lys Gln Arg Pro
 260 265 270
 Ala Ala Val Gln Thr Glu Glu Gln Tyr Arg Phe Leu Tyr His Thr Val
 275 280 285
 Ala Gln Met Phe Cys Ser Thr Leu Gln Asn Ala Ser Pro His Tyr Gln
 290 295 300
 Asn Ile Lys Glu Asn Cys Ala Pro Leu Tyr Asp Asp Ala Leu Phe Leu
 305 310 315 320

Arg Thr Pro Gln Ala Leu Leu Ala Ile Pro Arg Pro Pro Gly Gly Val
 325 330 335
 Leu Arg Ser Ile Ser Val Pro Gly Ser Pro Gly His Ala Met Ala Asp
 340 345 350
 Thr Tyr Ala Glu Glu Gln Lys Arg Gly Ala Pro Ala Gly Ala Gly Ser
 355 360 365
 Gly Thr Gln Thr Gly Thr Gly Thr Gly Ala Arg Ser Ala Glu Glu Ala
 370 375 380
 Pro Leu Tyr Ser Lys Val Thr Pro Arg Ala Gln Arg Pro Gly Ala His
 385 390 395 400
 Ala Glu Asp Ala Arg Gly Thr Leu Pro Gly Arg Val Pro Ala Asp Gln
 405 410 415
 Ser Pro Ala Gly Ser Gly Ala Tyr Glu Asp Val Ala Gly Gly Ala Gln
 420 425 430
 Thr Gly Gly Leu Gly Phe Asn Leu Arg Ile Gly Arg Pro Lys Gly Pro
 435 440 445
 Arg Asp Pro Pro Ala Glu Trp Thr Arg Val
 450 455

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 503 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

Met Glu Pro Ala Gly Pro Ala Pro Gly Arg Leu Gly Pro Leu Leu Cys
 1 5 10 15
 Leu Leu Leu Ala Ala Ser Cys Ala Trp Ser Gly Val Ala Gly Glu Glu
 20 25 30
 Glu Leu Gln Val Ile Gln Pro Asp Lys Ser Val Ser Val Ala Ala Gly
 35 40 45
 Glu Ser Ala Ile Leu His Cys Thr Val Thr Ser Leu Ile Pro Val Gly
 50 55 60
 Pro Ile Gln Trp Phe Arg Gly Ala Gly Pro Ala Arg Glu Leu Ile Tyr
 65 70 75 80
 Asn Gln Lys Glu Gly His Phe Pro Arg Val Thr Thr Val Ser Glu Ser
 85 90 95
 Thr Lys Arg Glu Asn Met Asp Phe Ser Ile Ser Ile Ser Asn Ile Thr
 100 105 110

Pro Ala Asp Ala Gly Thr Tyr Tyr Cys Val Lys Phe Arg Lys Gly Ser
 115 120 125
 Pro Asp Thr Glu Phe Lys Ser Gly Ala Gly Thr Glu Leu Ser Val Arg
 130 135 140
 Ala Lys Pro Ser Ala Pro Val Val Ser Gly Pro Ala Ala Arg Ala Thr
 145 150 155 160
 Pro Gln His Thr Val Ser Phe Thr Cys Glu Ser His Gly Phe Ser Pro
 165 170 175
 Arg Asp Ile Thr Leu Lys Trp Phe Lys Asn Gly Asn Glu Leu Ser Asp
 180 185 190
 Phe Gln Thr Asn Val Asp Pro Val Gly Glu Ser Val Ser Tyr Ser Ile
 195 200 205
 His Ser Thr Ala Lys Val Val Leu Thr Arg Glu Asp Val His Ser Gln
 210 215 220
 Val Ile Cys Glu Val Ala His Val Thr Leu Gln Gly Asp Pro Leu Arg
 225 230 235 240
 Gly Thr Ala Asn Leu Ser Glu Thr Ile Arg Val Pro Pro Thr Leu Glu
 245 250 255
 Val Thr Gln Gln Pro Val Arg Ala Glu Asn Gln Val Asn Val Thr Cys
 260 265 270
 Gln Val Arg Lys Phe Tyr Pro Gln Arg Leu Gln Leu Thr Trp Leu Glu
 275 280 285
 Asn Gly Asn Val Ser Arg Thr Glu Thr Ala Ser Thr Val Thr Glu Asn
 290 295 300
 Lys Asp Gly Thr Tyr Asn Trp Met Ser Trp Leu Leu Val Asn Val Ser
 305 310 315 320
 Ala His Arg Asp Asp Val Lys Leu Thr Cys Gln Val Glu His Asp Gly
 325 330 335
 Gln Pro Ala Val Ser Lys Ser His Asp Leu Lys Val Ser Ala His Pro
 340 345 350
 Lys Glu Gln Gly Ser Asn Thr Ala Ala Glu Asn Thr Gly Ser Asn Glu
 355 360 365
 Arg Asn Ile Tyr Ile Val Val Gly Val Val Cys Thr Leu Leu Val Ala
 370 375 380
 Leu Leu Met Ala Ala Leu Tyr Leu Val Arg Ile Arg Gln Lys Lys Ala
 385 390 395 400
 Gln Gly Ser Thr Ser Ser Thr Arg Leu His Glu Pro Glu Lys Asn Ala
 405 410 415
 Arg Glu Ile Thr Gln Asp Thr Asn Asp Ile Thr Tyr Ala Asp Leu Asn
 420 425 430

Leu Pro Lys Gly Lys Lys Pro Ala Pro Gln Ala Ala Glu Pro Asn Asn
 435 440 445
 His Thr Glu Tyr Ala Ser Ile Gln Thr Ser Pro Gln Pro Ala Ser Glu
 450 455 460
 Asp Thr Leu Thr Tyr Ala Asp Leu Asp Met Val His Leu Asn Arg Thr
 465 470 475 480
 Pro Lys Gln Pro Ala Pro Lys Pro Glu Pro Ser Phe Ser Glu Tyr Ala
 485 490 495
 Ser Val Gln Val Pro Arg Lys
 500

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 398 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

Met Pro Val Pro Ala Ser Trp Pro His Leu Pro Ser Pro Phe Leu Leu
 1 5 10 15
 Met Thr Leu Leu Leu Gly Arg Leu Thr Gly Val Ala Gly Glu Asp Glu
 20 25 30
 Leu Gln Val Ile Gln Pro Glu Lys Ser Val Ser Val Ala Ala Gly Glu
 35 40 45
 Ser Ala Thr Leu Arg Cys Ala Met Thr Ser Leu Ile Pro Val Gly Pro
 50 55 60
 Ile Met Trp Phe Arg Gly Ala Gly Ala Gly Arg Glu Leu Ile Tyr Asn
 65 70 75 80
 Gln Lys Glu Gly His Phe Pro Arg Val Thr Thr Val Ser Glu Leu Thr
 85 90 95
 Lys Arg Asn Asn Leu Asn Phe Ser Ile Ser Ile Ser Asn Ile Thr Pro
 100 105 110
 Ala Asp Ala Gly Thr Tyr Tyr Cys Val Lys Phe Arg Lys Gly Ser Pro
 115 120 125
 Asp Asp Val Glu Phe Lys Ser Gly Ala Gly Thr Glu Leu Ser Val Arg
 130 135 140
 Ala Lys Pro Ser Ala Pro Val Val Ser Gly Pro Ala Val Arg Ala Thr
 145 150 155 160
 Pro Glu His Thr Val Ser Phe Thr Cys Glu Ser His Gly Phe Ser Pro
 165 170 175

Arg Asp Ile Thr Leu Lys Trp Phe Lys Asn Gly Asn Glu Leu Ser Asp
 180 185 190
 Phe Gln Thr Asn Val Asp Pro Ala Gly Asp Ser Val Ser Tyr Ser Ile
 195 200 205
 His Ser Thr Ala Arg Val Val Leu Thr Arg Gly Asp Val His Ser Gln
 210 215 220
 Val Ile Cys Glu Met Ala His Ile Thr Leu Gln Gly Asp Pro Leu Arg
 225 230 235 240
 Gly Thr Ala Asn Leu Ser Glu Ala Ile Arg Val Pro Pro Thr Leu Glu
 245 250 255
 Val Thr Gln Gln Pro Met Arg Ala Glu Asn Gln Ala Asn Val Thr Cys
 260 265 270
 Gln Val Ser Asn Phe Tyr Pro Arg Gly Leu Gln Leu Thr Trp Leu Glu
 275 280 285
 Asn Gly Asn Val Ser Arg Thr Glu Thr Ala Ser Thr Leu Thr Glu Asn
 290 295 300
 Lys Asp Gly Thr Tyr Asn Trp Met Ser Trp Leu Leu Val Asn Thr Cys
 305 310 315 320
 Ala His Arg Asp Asp Val Val Leu Thr Cys Gln Val Glu His Asp Gly
 325 330 335
 Gln Gln Ala Val Ser Lys Ser Tyr Ala Leu Glu Ile Ser Ala His Gln
 340 345 350
 Lys Glu His Gly Ser Asp Ile Thr His Glu Pro Ala Leu Ala Pro Thr
 355 360 365
 Ala Pro Leu Leu Val Ala Leu Leu Leu Gly Pro Lys Leu Leu Leu Val
 370 375 380
 Val Gly Val Ser Ala Ile Tyr Ile Cys Trp Lys Gln Lys Ala
 385 390 395